DOCUMENT RESUME

ED 371 580 FL 022 051

AUTHOR	Davies, Alan
TITLE	Simply Defining: Constructing a Dictionary of Language Testing.
PUB DATE	93
NOTE	15p.; In: Tickoo, M. L., Ed. Simplification: Theory and Application. Anthology Series 31; see FL 022 043.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Definitions; *Dictionaries; Educational Media; *Encyclopedias; Foreign Countries; *Graduate Students; Higher Education; *Language Attitudes; Language Research; *Language Tests; Language Usage; *Readability; Testing
IDENTIFIERS	Simplification (Language); University of Melbourne (Australia)

ABSTRACT

This paper examines the nature of dictionaries and encyclopedias, focusing on some of the preparation that has gone into the construction of a dictionary of language testing at the University of Melbourne (Australia). It discusses the purpose of such dictionaries, the nature and size of dictionary/encyclopedia entries, and the readability of entries. It also reports on an experiment to determine the proper length and difficulty level of possible entries. Twenty-one M.A. students were asked to read three sets of entries and comment on their length, difficulty, and lexical density. It was found that there was substantial agreement among the students as to whether several entries were too long, too short, or about right, and if they were difficult, easy, or about right. An appendix contains copies of the possible dictionary entries. (MDM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 371 580

SIMPLY DEFINING CONSTRUCTING A DICTIONARY OF LANGUAGE TESTING

Alan Davies

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wong Kim

Wong

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL022051

SIMPLY DEFINING CONSTRUCTING A DICTIONARY OF LANGUAGE TESTING

Alan Davies

'He that undertakes to compile a Dictionary undertakes that which, if it comprehends the full extent of his design, he knows himself unable to perform'

(Johnson 1773)

C S Lewis (1960) comments on the term 'simple' thus: 'Every compound, or so we hope, can in principle be resolved into *simple* ingredients, ingredients which are internally homogeneous. And as the compound is a compound, so these ultimate ingredients are *simples*.' (166). 'What is simple or plain is the reverse of complicated. A complicated process is hard to learn and a complicated argument hard to follow. Therefore simple comes to mean 'easy'. The idea that it is within the capacity of those who are simple (in the sense 'unskilled') may perhaps have helped this development.' (174). 'I describe the final state of the word as a semantic sediment. What effectively remains is not this or that precise sense but a general appealingness or disarmingness.' (179).

Disarming indeed! Does this mean that there is no objective judgement as to what is simple or difficult, that it is all a question of appealingness, a kind of political correctness (P.C.) of the lexicon?

As far as texts are concerned there has been a great deal of investigation of what makes for ease of readability. Jeanne Chall (1984) summarises the findings of this work, thus:

'What makes text easy or hard to read and comprehend? The research in readability has uncovered over one hundred factors related to difficulty - such factors as vocabulary, sentences, ideas, concepts, text organization, content, abstractness, appeal, format and illustrations. Of these factors, the two found consistently to be most strongly associated with comprehensibility are vocabulary difficulty and sentence length. Various forms of these two factors are included in most of the currently used readability formulas. The strongest factor of the two is vocabulary difficulty - measured either by a count of unfamiliar words, hard words, words of low frequency, words of three or more syllables, or words of 7 letters or

BEST COPY AVAILABLE

more. All word measures are highly interrelated. Once a vocabulary factor is used in a formula, another adds little to the prediction. Average sentence length is the second strongest and second most widely used measure of difficulty in readability formulas. It is very highly related to other measures of syntactic difficulty, and therefore only one sentence factor is usually used in a formula. It is also substantially associated with vocabulary difficulty. A vocabulary and a sentence factor together predict the comprehension difficulty of written text to a high degree of accuracy. The multiple correlations run about 0.7 to 0.9 with reading comprehension in multiple-choice or cloze tests.' (237-8).

Chall does not of course deny that there are other factors that contribute to text readability but for the purposes of this paper we will focus on the two she details, vocabulary and sentence length, although we will not here use any readability formulas.

Our concern is with the preparation of information for a dictionary (or more specifically for an encyclopedic dictionary, see below for the distinction). The paper reports some of the preparation that has gone into the construction of a Dictionary of Language Testing (Davies 1992) at the NLLIA Language Testing Centre, University of Melbourne.

Making information available to others, the institutionalised preparation of data to inform, the heart of all pedagogy, is necessarily a process of simplification. Or to put it another way we may define simplification as the pedagogic delivery of information. Not *simple*, that is an issue of possible difference existing in nature whereby for example an X is more difficult than a Y. Observe that there is no limitation here as to context; in other words the simple-difficult continuum in nature is, we suggest, an absolute one. Because of course there may always be a relative difference of awareness in that what is difficult for learners (for example children) may be simple for those who are advanced (for example adults).

The history of readability research records continuing attempts to bring closer the text and the reader. Readability in other words is seen to be a function of their interaction so that measures of difficulty for texts and of comprehension for readers are both necessary. Taylor's development of the cloze technique (1957) was a deliberate attempt to combine these two variables in one measure which could be used to ascertain readability of newspapers. Davies (1984), focusing on the second language situation, reports on an experiment in which an original text and its simplified version were compared in terms of their comprehension by a group of Japanese teachers of English. The hypothesis was that the linguistically simpler text would be comprehended better than the original text. The hypothesis

was supported. In that experiment the measure of comprehension was verbatim cloze.

In our current task of writing a dictionary, care must obviously be taken that the definition or explanations can be understood by the reader. That means targeting the dictionary carefully so that the explanations have appropriate readability for their audience. Dictionaries therefore have a built-in pedagogic function. This raises the questions of just what a dictionary is and in particular what sort of word-book is needed for a professional-academic audience. Some views by dictionary makers will be of interest:

Abercrombie N., Hill S., Turner B S (1984) claim that 'A dictionary of sociology is not just a collection of definitions, but inevitably a statement of what the discipline is. It is also prescriptive in suggesting lines of development and consolidation. The problem of definition in a subject as diverse and dynamic as sociology is to strike a balance between an existing consensus, however fragile and temporary, and a developing potential. The unifying theme of this dictionary is our conviction that sociology is an autonomous, elaborated and vital discipline within the social science corpus. Our enthusiasm for the subject was sustained rather than diminished by the experience of seeking precision within the conflicting range of perspectives that constitute modern sociology.' (p. vii). 'A statement of what the discipline is': a tall order indeed but nevertheless inevitably what all dictionary making assumes in its normative role.

West and Endicott (1935/59) maintain 'This English Dictionary is written specially for the foreigner. It explains to him in words 'which he knows the meaning of words and idioms which he does not know.' (p. iii). In words which s/he knows the meaning of, stresses the welcomed linguistic straitjacket of dictionary making.

Angeles (1981) states that his dictionary 'is intended as an at-hand reference for students, laypersons, and teachers. It can be used as a supplement to texts and philosophy readings; it can also be consulted for philosophy's own enjoyment and enlightenment.' (p. ix). Even the 'laypersons' Angeles refers to must surely be informed, interested, educated and so on. Audience is critical and when it includes students necessarily demands some measure of simplification, if not of language, certainly of substance.

The idea for writing a Dictionary of Language Testing by colleagues in the NLLIA Language Testing Centre, University of Melbourne arose out of several needs; the need for a kind of in-house set of glosses so that we all know what we are talking about, our own register; then, as with text-books, the necessary compromise between the profession (those like ourselves working in Applied Linguistics and language testing) and the public. Such a compromise targets those with general rather than specialist knowledge, ie MA students of the relevant disciplines. Very much, in fact, like the audience targeted by Richards, Platt and Weber in their Longman Dictionary of Applied Linguistics (1985). We have found that working together on this dictionary is a felicitous way of sharing and educating one another, precisely because it defines, explores and creates the very register we need for our work. Those working on the dictionary (Alan Davies, Tim McNamara, Cathie Elder, Annie Brown, Tom Lumley, Chris Corbel, Yap Soon Hock), are all very much part-time and we are aware that this is a long term task. The Longman Dictionary of Applied Linguistics, after all, was 4 years in the making. A larger contributing group of authors provides wide coverage but needs marshalling and organising. And in such an exercise organisation is essential.

As I have already suggested the problems have to do with audience and definition; but equally important are selection and coverage, scope and format of entry. In attempting to reach agreement over these matters we have been helped by the realisation that such concerns are not at all new in lexicography. If indeed what we are doing is lexicography.

Let me quote from Kipfer's Workbook (1984:1) 'A dictionary is a reference book containing the words of a language or language variety, usually alphabetically arranged, with information on their forms, pronunciations, functions, meanings and idiomatic uses. A dictionary may be more than a reference book about words; it can contain biographical and geographical knowledge as well as lists of colleges and universities, weights and measures, and symbols; and the introduction may include articles about aspects of language and dictionary making. The entries themselves may contain not only pronunciations and meanings, but information about grammar and usage and even the kind of information an encyclopedia gives about the thing the word names.'

There are indeed many terms for our type of ambition: are we writing a dictionary, an encyclopedia, a word list, a glossary, a reference list? Opitz writes of a 'segmental dictionary', but is that what it is or a glossary, perhaps -ie a list of technical terms rather than an attempt 'to isolate a distinct register' which is what Opitz means by a segmental dictionary. (Opitz 1983: 58).

But is it a glossary? Hartmann defines glossary as a 'word-list with explanation of meanings' (Hartmann 1983: 223). Moulin describes a glossary as a list of glosses appended to text, often specialised, and details two techniques of ordering, by areas of interest and by alphabet: 'most authors (of specialist dictionaries) are neither linguists nor professional lexicographers, but specialists in the particular discipline...these glossaries are commissioned...to try and introduce a measure of normalisation in the use of specialist terms and thus facilitate the exchange of information' (Moulin 1983: 146). We will return to that concern for a 'measure of normalisation'.

Is it an encyclopedia? Hartmann tells us that encyclopedic information has to do with 'practical knowledge of things versus lexical information' (Hartmann 1983: 223). A more elaborate distinction is made by Read (1976: 713ff) quoted in McArthur 1986: 'The distinction between a dictionary and an encyclopedia is easy to state but difficult to carry out in a practical way: a dictionary explains words, whereas an encyclopedia explains things. Because words achieve their usefulness by referring to things, however, it is difficult to construct a dictionary without considerable attention to the objects and abstractions designated'.

McArthur reminds us that the Encyclopedia Britannica had its origin in Edinburgh. Notice its original title: 'The Encyclopedia Britannica or a Dictionary of Arts and Sciences, compiled upon a New Plan' (Edinburgh 1768-71, sponsored by the Society of Gentlemen in Scotland). The Britannica was a very obvious product (no doubt influenced by the French philosophers) of the Scottish Enlightenment, that high point in Scottish history, when Scotland truly was the clever country. From that high point we are brought down to earth by the comment of William Smellie, one of the original authors: 'with pastepot and scissors I composed it' (W. Smellie in Kogan 1956: 14, quoted in McArthur: 106-7).

McArthur suggests as a way of resolving the overlap in the uses of the terms Dictionary and Encyclopedia that it is probably best not to bother. 'The simplest way' he says 'of resolving the tension seems to be to accept the way in which the early encyclopedists handled the matter. In this dilemma we in fact work along a continuum rather than within separate containers, where one extreme is words and words alone, and the other is referents and referents alone' (McArthur 1986: 104). At one end of McArthur's continuum is the dictionary, at the other the encyclopedia and in between the encyclopedic dictionary.

McArthur suggests as a way of relating dictionaries and encyclopedias (which in the USA and France, but not in the UK, have, he says always been linked) the following pair of terms 'that could be useful in studying the world of reference materials:

1. micro-lexicography, which deals with the world of words and the wordbook proper (which in most instances is an alphabetic dictionary).
2. macro-lexicography, which shades out into the world of things and subjects, and centres on compendia of knowledge (which in most instances are encyclopedias, which in most instances nowadays are also alphabetic') (McArthur 1986: 109).

When we look at attempts within the field of Applied Linguistics two well-known products are the Longman Dictionary of Applied Linguistics and the First Dictionary of Linguistics and Phonetics. In the Introduction to their Longman Dictionary of Applied Linguistics, Richards, Platt and Weber (1985) ask who is this dictionary intended for and conclude that it is intended for students of Applied Linguistics and General Linguistics; language teachers both in training and in the field. (1985: v).

'Our aim' they say 'has been to produce clear and simple definitions which communicate the basic and essential meanings of a term in non-technical language. Definitions are self-contained as far as possible, but cross references show links to other terms and concepts, and references provide information where a fuller discussion of a term or concept can be found.' (p. vii).

More helpful to us in our grappling with the problems is Crystal who in the Preface to his A First Dictionary of Linguistics and Phonetics (1980) confesses 'I remain doubtful even now whether the most appropriate title for this book is "dictionary". The definitional parts of the entries by themselves were less illuminating than one might have expected; and consequently it proved necessary to introduce in addition a more discursive approach, with several illustrations, to capture the significance of a term. Most entries accordingly contain encyclopedic information about such matters as the historical context in which a term was used, or the relationship between a term and others from associated fields' (p. 5).

'Each entry', Crystal continues, 'is self-contained: that is there are no obligatory cross-references to other entries to complete the exposition of a sense. Nor have I made use of the convention 'See Y' after looking up a term...I have preferred to work on the principle that, as most dictionary-users open a dictionary with a single problematic term in mind, they should be given a satisfactory account of that term as immediately as possible. I therefore explain competence under COMPETENCE, performance under PERFORMANCE, and so on. As a consequence of the interdependence of these terms, however, this procedure means that there must be some repetition: at least the salient characteristics of the term performance must be incorporated into the entry for COMPETENCE, and vice-

versa. This repetition would be a weakness, if the book were read from cover to cover; but a dictionary should not be used as a text-book, and while the result has been a somewhat longer volume than would have been the case if the 'See...' convention had been used, I remain convinced of the greater benefits of look-up convenience and entry coherence.' (p. 5).

After some preliminary trials, pilot entry writing and a small-scale survey of the entries among teachers and MA students, we determined on the following guidelines for ourselves:

1. the entries should be on the encyclopedia side of McArthur's continuum, more than language-definitional, explaining where appropriate;
2. they should where possible (and appropriate) give examples so as to situate the explanation;
3. they should accept overlap, in Crystal's sense, so that referring to other entries for necessary explanation would be avoided, except where necessary for informative purposes; citations would be minimised except in the sense of the informative purpose above;

where possible one clear definition should be attempted, in other words coming down on the side of being normative rather than descriptive. We have taken the view that unlike a truly descriptive dictionary (such as the OED) it is our role to contain and confine to 'try and introduce a measure of normalisation in the use of specialist terms and thus facilitate the exchange of information.' (Moulin 1983: 146).

Whether what we are doing therefore should be called a dictionary or an encyclopedia is really beside the point. But while it does veer towards the encyclopedia side of the McArthur continuum it retains important aspects of dictionary-ness. It does attempt definitions, it avoids essays (so it is not a Glossary either: 'I have retained the procedure of organizing the Glossary as a series of essays' (Abrams 1981: v) but unlike many dictionaries it has no information of a pronunciation kind (though obviously it would not eschew this where it seemed relevant) nor does it systematically contain historical material about derivations. So it probably is what McArthur calls an encyclopedic dictionary.

We are in our Language Testing Dictionary concerned to establish a uniform style of entry and at the same time to ensure adequate coverage. To illustrate these questions and through them the importance in our view of being more encyclopedic than dictionary-like, I turn now to a comparison of alternative entries.

We are planning for about four hundred entries in our dictionary. Given the choice between the A version and the B version below, our present view is very much in favour of the B version, even though use of the A version would permit a larger number of entries. In each case the A version is much shorter than the B version, in some sense therefore the B version is more encyclopedic like and the A version more dictionary like. In making the comparisons reported below our hypothesis was that because of our perceived nature of this Dictionary the B versions were more likely to be readable than the A versions.

Experiment and Results

A class of MA students (N = 21) were asked to read three sets of entries (see Appendix) and comment on (1) their length - were they too long, too short or about right; and (2) their difficulty - were they too difficult, too easy or about right. With hindsight it is apparent that these were unsatisfactory choices to have to make. What after all does 'too easy' mean? Nevertheless the responses do provide us with some indication of the readability of the contrasting versions we had provided.

Next a comparison was made between the (a) and (b) versions on the basis of their lexical density (Halliday 1985). Lexical density is an indication of the ratio of lexical to grammatical loading clause by clause. Halliday reports that in informal spoken English lexical density is about 2; in adult written language it is typically more dense, say about 6 per clause. In scientific writing it can be as high as 10-13 per clause. That is one reason why scientific writing is often so difficult except to the expert. It is also an explanation of why newspaper headlines can be almost uninterpretable, unless you know exactly what is currently at issue.

Here are the summed responses of the Masters' students alongside the lexical density finding for each entry.

Draft Entries

	1A	1B	2A	2B	$3A_1$	$3A_2$	3B
Too long	0	9	0	7	0	0	16
Too short	14	0	17	0	19	8	0
About right	7	12	3	14	2	13	4
Difficult	5	10	6	3	6	2	9
Easy	7	1	9	0	13	4	1
About right	9	10	4	18	2	14	9
Lex Density	7	7	12	5.8	11	6.5	6.8

Table 1. Draft Entries : Responses and Lexical Density

As a rule of thumb we suggest that 50+% approval for an entry indicates acceptability. On that basis, two entries, 2B and $3A_2$ may be labelled acceptable. That judgement is supported by the lexical density comparison for these two entries. Note that 2B has a much lower lexical density than 2A (less than half) and that $3A_2$, while much lower on lexical density than $3A_1$, is marginally lower than 3B. The problem with 3B, which we had predicted would be rated as more acceptable by the class, seems to be sheer length. It is interesting that (see Table) there is substantial agreement (16/21) that this entry is too long and yet at the same time as many as 9/21 accorded it 'about right' for difficulty. A similar result emerges for Entry 1 where there is no separation between the A and the B versions in terms of difficulty (both entries obtain a lexical density result of 7). At the same time (and here is the similarity with entry 3B), there were 9 responses in the 'about right' response for 1A and 10 for 1B. We might therefore suggest that acceptability as indicated by being accorded 'about right' for difficulty is in part a function of lexical density. Where lexical density does not discriminate (as in the 1A and 1B entries) neither choice is regarded as being acceptable. Where there may be little to choose in terms of lexical density (as between Entries $3A_2$ and 3B) length of an entry may militate against the choice of an entry (as with Entry 3b).

Of course it may be queried whether a response of 'about right' is appropriate, whether indeed (as with the figure quoted above for scientific writing, a lexical density of between 10 and 13) specialists tolerate a high density, a 'more difficult' entry. But that is after all a comment on the sampling of our responses and it is indeed our contention that the class whose responses we report here are the

appropriate audience for our Dictionary of Language Testing, students on Masters courses who are in the process of being introduced to classes in Applied Linguistics, including Language Testing. Specialists in the field (pace Abercrombie et al 1984) are not our concern in this task; to what extent a dictionary for specialists is a viable activity we remain unclear about. In our view dictionaries (and here we would agree with Abercrombie et al 1984) are always normative in the sense that they are drawing boundaries with a pedagogic intent. No dictionary, certainly not a segmental dictionary, can ever satisfy the specialist! For that is what we have decided to term our effusion, a segmental dictionary: as such it retains its professional/vocational/registral association and at the same time its normative/pedagogical purpose which is what makes our attempts at simplification necessary.

'I am not so lost in lexicography as to forget that words are the daughters of earth, and that things (including deeds) are the sons of heaven' (Samuel Johnson, in the Preface of his Dictionary 1755)

REFERENCES

ABERCROMBIE N, Hill S, and Turner B S (eds) (1984). The Penguin Dictionary of Sociology. Penguin, London.

ABRAMS M H (1981). A Glossary of Literary Terms. (4th ed.) Holt-Saunder, Japan.

ANGELES P A (1981). Dictionary of Philosophy. Barnes and Noble/Harper and Row, New York.

BACHMAN L F (1990). Fundamental Considerations in Language Testing. Oxford University Press, Oxford.

CHALL JEANNE S (1984). Readabilty and Prose Comprehension: Continuities and Discontinuities: Understanding Reading Comprehension. (ed). Flood J International Reading Association, Newark, Delaware. 233-246.

CRYSTAL D (1980). A First Dictionary of Linguistics and Phonetics. Andre Deutsch, London.

DAVIES A (1984). 'Simple, Simplified and Simplification - What is Authentic?'. In Alderson A C and Urquhart A H. Reading in a Foreign Language. Longman, London 181-195.

DAVIES A (February/March 1992). *'The Role of the Segmental Dictionary in Professional Validation' paper given at the 14th Annual Language Testing Research Colloquium, University of British Columbia.*

HALLIDAY M A K (1985). *Spoken and Written Language. Deakin University Press, Deakin University, Victoria 3217, Australia.*

HARTMANN R R K (ed.) (1983). *Lexicography: Principles and Practice. Academic Press, London.*

JENSEN A R (1980). *Bias in Mental Testing. The Free Press, New York.*

JOHNSON S (1755). *The Dictionary, with a Grammar and History of the English Language. London, Preface to 4th Ed. 1773.*

KIPFER B (1984). *Workbook of Lexicography. Exeter Linguistics Series, Exeter University.*

KOGAN H (1958). *The Great E.B., The Story of the Encyclopedia Britannica. Chicago University Press, Chicago.*

LEWIS C S (1960). *Chapter 7 'Simple' in Studies in Words. CUP.*

MCARTHUR TOM (1986). *Worlds of Reference: Lexicography, Learning and Language from the Clay Tablet to the Computer. Cambridge University Press, Cambridge.*

MOULIN A (1983). *'LSP Dictionaries for EFL Learners'. In Hartmann (op. cit) 144-152.*

OPITZ K (1983). *'On Dictionaries for Special Registers'. In Hartmann (op. cit) 53-64.*

READ A W (1976). *'Dictionaries' Encyclopedia Britannica. 15th Edition, New York.*

RICHARDS J, Platt J, and Weber H (1985). *Longman Dictionary of Applied Linguistics. Longman, London.*

TAYLOR W L (1953). *'"Cloze Procedure": A New Tool for Measuring Readability'. Journalism Quarterly 30: 415-433.*

WEST M P, and Endicott J G (1935/59). *The New Method English Dictionary. Longman, London.*

APPENDIX

Entry 1A
Variance: (in testing and statistics) a measure of the **DISPERSION** of a **SAMPLE**. The variance of a set of scores, on a test for example, would be based on how much the scores obtained differ from the **MEAN**, and is itself the square of the **STANDARD DEVIATION**.

Entry 1B
Variance: a statistical measure of the **DISPERSION** of a **SAMPLE** which can be expressed in a standardised square root form as a **STANDARD DEVIATION**, that is to say that the variance of a sample is the standard deviation squared. The dispersion of a sample on one measure (or test) may be compared with its dispersion on another; this comparison is referred to as the shared variance. Such a comparison is achieved by means of a **CORRELATION** and further comparisons of dispersion on other measures are carried out by means of **ANALYSIS OF VARIANCE**. The square of the correlation indicates in percentage terms the shared variance between two measures. Two tests which correlate 0.7 would have a shared variance of 49%, while a higher correlation of 0.9 would still indicate a shared variance of only 81%, leaving 19% of the variance unexplained by the overlap between the two tests.

Entry 2A
Analysis of Variance: a statistical procedure used for estimating the relative effects of different sources of variance on test scores (ANOVA) (Bachman 1990:193).

Entry 2B
Analysis of Variance: a statistical procedure which combines correlations of several variables with one another and against a common criterion, with the intention of determining the influence (if any) of one variable upon another. Analysis of Variance (or Anova as it is often called) helps observers to avoid simplistic conclusions assuming causality between one variable and a criterion. Anova is commonly available on computer statistical packages. Example: success at the end of an intermediate language course is shown to be significantly correlated with scores on an entry language test; when two other variables, age and motivation, are added to the study, it might turn out that entry scores no longer predict or do so only in relation to age and/or motivation; or that age is now so important a predictor that, when the 'variance' due to age is removed from the analysis, what remains for entry scores is trivial. (see: variance, correlation, criterion, variable, predict).

Entry $3A_1$

Bias: 'systematic error associated with any type of group membership, sex and age group membership included.' (Jensen 1980).

Entry $3A_2$

Bias: 'systematic differences in test performance that are the result of differences in individual characteristics, other than the ability being tested, of test takers.' (Bachman 1990: 271).

Entry 3B

Bias: bias is defined by Jensen (1980) as 'systematic error associated with any type of group membership, sex and age group membership included'. The terms *systematic error* and *group* are important in this definition. In a trivial way all tests are biased against individuals who lack knowledge or skill. But since that is what tests are designed to do, such 'bias' or, better, discrimination is not systematic error, that is to say there will be random error as in all measurement but it is not systematic or deliberate. The *group* issue is more problematic in situations of norm conflict such as recent migrant communities. Should children from such communities with only a few years of schooling in the target language (eg English) take the same language tests as first language speaking children? In conflict are (1) the general educational norms and standards of the host community and (2) what it is reasonable to expect in terms of English language proficiency of the migrant children. In situations where the first consideration weighs more heavily, migrant children will take the same English test as first language children. In situations where the second consideration is more important, a more specialised test of ESL may be used.